REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

AD-A204 930

1a.	
1b. RESTRICTIVE MARKINGS	
2a.	
2b.	
3. DISTRIBUTION/AVAILABILITY OF REPORT	Approved for Public Release; Distribution Unlimited
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	Technical Report No. 249
5. MONITORING ORGANIZATION REPORT NUMBER(S)	AFOSR-TR- 89-0120
6a. NAME OF PERFORMING ORGANIZATION	University of North Carolina Center for Stochastic Processes
6b. OFFICE SYMBOL (If applicable)	
7a. NAME OF MONITORING ORGANIZATION	AFOSR/NM
6c. ADDRESS (City, State, and ZIP Code)	Statistics Department CB #3260, Phillips Hall Chapel Hill, NC 27599-3260
7b. ADDRESS (City, State, and ZIP Code)	Bldg. 410 Bolling Air Force Base, DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION	AFOSR
8b. OFFICE SYMBOL (If applicable)	NM
9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER	F49620 85C 0144
8c. ADDRESS (City, State, and ZIP Code)	Bldg. 410 Bolling AFB, DC 20332-6448

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
6.1102F	2304	A5	

11. TITLE (Include Security Classification)
Bias and variance approximations for estimators of extreme quantiles

12. PERSONAL AUTHOR(S)
Smith, R.L.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Preprint	FROM Nov 88 TO Nov 89	1988, Nov.	40

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Extreme values; generalised extreme value distribution, generalised Pareto Distribution, Gumbel Distribution, maximum likelihood, threshold methods.
XXXXX	XXXXXXXXXX	XXX	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Most techniques for estimating extreme values are based on the assumption of a parametric family motivated by extreme value limit theory. This creates two sources of estimation error: the ordinary estimation variance and a bias created by misspecification of the parametric model. In this paper approximate formulae are derived for the bias and variance of four widely studied estimators. This allows comparison among the different estimators. The development relies on recent work on probabilistic approximations in extreme value theory. Key words.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
Major Brian W. Woodruff	(202)767-5026	AFOSR/NM

DD Form 1473, JUN 86 — Previous editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFED

CENTER FOR STOCHASTIC PROCESSES

Department of Statistics
University of North Carolina
Chapel Hill, North Carolina

BIAS AND VARIANCE APPROXIMATIONS FOR

ESTIMATORS OF EXTREME QUANTILES

by

Richard L. Smith

Technical Report No. 249

November 1988

179. R. Brigola, Remark on the multiple Wiener integral, Mar. 87.

180. R. Brigola, Stochastic filtering solutions for ill-posed linear problems and their extension to measurable transformations, Mar. 87.

181. G. Samorodnitsky, Maxima of symmetric stable processes, Mar. 87.

182. H.L. Hurd, Representation of harmonizable periodically correlated processes and their covariance, Apr. 87. *J. Multivariate Anal.*, to appear.

183. H.L. Hurd, Nonparametric time series analysis for periodically correlated processes, Apr. 87. *IEEE Trans. Information Theory*, to appear.

184. T. Mori and H. Oodaira, Freidlin-Wentzell estimates and the law of the iterated logarithm for a class of stochastic processes related to symmetric statistics, May 87. *Yokohama Math. J.*, to appear.

185. R.F. Serfozo, Point processes, May 87. *Operations Research Handbook on Stochastic Processes*, to appear.

186. Z.D. Bai, W.Q. Liang and W. Vervaat, Strong representation of weak convergence, June 87.

187. O. Kallenberg, Decoupling identities and predictable transformations in exchangeability, June, 87.

188. O. Kallenberg, An elementary approach to the Daniell-Kolmogorov theorem and some related results, June 87. *Math. Nachr.*, to appear.

189. G. Samorodnitsky, Extrema of skewed stable processes, June 87. *Stochastic Proc. Appl.*, to appear.

190. D. Nualart, M. Sanz and M. Zakai, On the relations between increasing functions associated with two-parameter continuous martingales, June 87.

191. F. Avram and M. Taqqu, Weak convergence of sums of moving averages in the α-stable domain of attraction, June 87.

192. M.R. Leadbetter, Harald Cramér (1893-1985), July 87. *ISI Review*, 56, 1988, 89-97.

193. R. LePage, Predicting transforms of stable noise, July 87.

194. R. LePage and B.M. Schreiber, Strategies based on maximizing expected log, July 87.

195. J. Rosinski, Series representations of infinitely divisible random vectors and a generalized shot noise in Banach spaces, July 87. *Ann. Probability*, to appear.

196. J. Szulga, On hypercontractivity of α-stable random variables, $0<\alpha<2$. July 87.

197. I. Kuznezova-Sholpo and S.T. Rachev, Explicit solutions of moment problems I, July 87. *Probability Math. Statist.*, 10, 1989, to appear.

198. T. Hsing, On the extreme order statistics for a stationary sequence, July 87. *Stochastic Proc. Appl.* 29, 1988, 155-169.

199. T. Hsing, Characterization of certain point processes, Aug. 87. *Stochastic Proc. Appl.* 26, 1987, 297-316.

200. J.P. Nolan, Continuity of symmetric stable processes, Aug. 87. *J. Multivariate Anal.*, 1988, to appear.

201. M. Marques and S. Cambanis, Admissible and singular translates of stable processes, Aug. 88.

202. O. Kallenberg, One-dimensional uniqueness and convergence criteria for exchangeable processes, Aug. 87. *Stochastic Proc. Appl.* 28, 1988, 159-183.

203. R.J. Adler, S. Cambanis and G. Samorodnitsky, On stable Markov processes, Sept. 87.

204. G. Kallianpur and V. Perez-Abreu, Stochastic evolution equations driven by nuclear space valued martingales, Sept. 87. Appl. Math. Optimization, 17, 1988, 237-272.

205. R.L. Smith, Approximations in extreme value theory, Sept. 87.

206. E. Willekens, Estimation of convolution tails, Sept. 87.

207. J. Rosinski, On path properties of certain infinitely divisible processes, Sept. 87. Stochastic Proc. Appl., to appear.

208. A.H. Korezlioglu, Computation of filters by sampling and quantization, Sept. 87.

209. J. Bather, Stopping rules and observed significance levels, Sept. 87. *Sequential Anal.*, to appear.

210. S.T. Rachev and *J.E. Yukich*, *Convolution metrics and rates of convergence in the central limit theorem*, Sept. 87. *Ann. Probability*, 1988, to appear.

211. M. Fujisaki, Normed Bellman equation with degenerate diffusion coefficients and its applications to differential equations, Oct. 87.

212. G. Simons, Y.C. Yao and X. Wu, Sequential tests for the drift of a Wiener process with a smooth prior, and the heat equation, Oct. 87.

213. R.L. Smith, Extreme value theory for dependent sequences via the Stein-Chen method of Poisson approximation, Oct. 87.

214. C. Houdré, A vector bimeasure integral with some applications, June 88 (Revised).

215. M.R. Leadbetter, On the exceedance random measures for stationary processes, Nov. 87.

216. M. Marques, A study on Lebesgue decomposition of measures induced by stable processes, Nov. 87 (*Dissertation*).

217. M.T. Alpuim, High level exceedances in stationary sequences with extremal index, Dec. 87. Stochastic Proc. Appl., to appear.

218. R.F. Serfozo, Poisson functionals of Markov processes and queueing networks, Dec. 87.

219. J. Bather, Stopping rules and ordered families of distributions, Dec. 87. *Sequential Anal.*, 7, 1988, 111-126.

220. S. Cambanis and M. Maejima, Two classes of self-similar stable processes with stationary increments, Jan. 88.

221. H.P. Hucke, G. Kallianpur and R.L. Karandikar, Smoothness properties of the conditional expectation in finitely additive white noise filtering, Jan. 88. *J. Multivariate Anal.*, to appear.

BIAS AND VARIANCE APPROXIMATIONS FOR ESTIMATORS OF EXTREME QUANTILES

Richard L. Smith

Department of Mathematics

University of Surrey

Guildford GU2 5XH, England.

Summary

Most techniques for estimating extreme values are based on the assumption of a parametric family motivated by extreme value limit theory. This creates two sources of estimation error: the ordinary estimation variance and a bias created by misspecification of the parametric model. In this paper approximate formulae are derived for the bias and variance of four widely studied estimators. This allows comparison among the different estimators. The development relies on recent work on probabilistic approximations in extreme value theory.

Keywords Extreme values, Generalised Extreme Value Distribution, Generalised Pareto Distribution, Gumbel Distribution, maximum likelihood, threshold methods.

September 1988.

1. INTRODUCTION

Two principal methods have been developed for inference about extreme quantiles or tails of distributions. The older is based on extreme value distributions for subsample maxima. Statistical aspects of this have been developed by Prescott and Walden (1980), Hosking (1984) and Smith (1985), amongst others. The other method is based on the Generalised Pareto distribution for exceedances over high thresholds. This was introduced by Pickands (1975) and developed by Davison (1984), Smith (1984) and Hosking and Wallis (1987), though there is a much larger literature of related techniques in hydrology (NERC 1975).

A major issue with both of these methods is that certain limiting distributions are used as statistical models. These limiting distributions are not exact in finite samples, and so there are two sources of error. One source is the usual variance of estimators of the model. The other source is a bias created by the fact that the assumed model is not exact. This creates a bias versus variance conflict of the kind familiar from density estimation, nonparametric regression, and other statistical techniques involving smoothing.

Previous studies of this feature have been Davis and Resnick (1984) and Smith (1987) on threshold methods and Cohen (1987, 1988) on classical extreme value methods. Joe's (1987) results partly unify the two methods. There is a long literature on estimating an index of regular variation, of which Hall (1982) and Hall and Welsh (1984, 1985) have done most to make the connection with smoothing problems. Recent references on this include Csorgo, Horvath and Revesz (1987), Reiss (1987) and Bierlant and Teugels (1987). Simulation results have been given by Boos (1984), Gomes (1986) and Joe (1987).

One reason why the literature on this topic has been so disjointed is that, although there has been much literature on rates of convergence in extreme value theory, no unified approach has emerged. In particular, the three domains of attraction tend to be treated entirely separately, so that the literature is three times as long as it should be. Recently, however, Smith (1988) has proposed a new approach bringing together the three domains of attraction. In the present paper, my aim is to show how this new approach leads to explicit (approximate) expressions for the bias and variance of estimators of extreme quantiles. The principal features which distinguish this approach from its predecessors are that the three domains of attraction are dealt with together, and that the method applies to a very wide variety of estimators. To illustrate this, four particular estimators are studied in detail. The results may therefore be used to compare one estimation method with another, as well as providing a theoretical resolution of such questions as choosing the best threshold. They may also suggest automatic or adaptive techniques of, for example, choosing the threshold, but I do not consider this aspect in any detail. For this reason the results should be viewed as providing general qualitative guidelines rather than determining a specific procedure for a particular set of data.

The organisation of the paper is as follows. Section 2 reviews the development of Smith (1988) on probabilistic approximations to extreme value distributions. Sections 3-6 develop, in turn, bias and variance approximation for four previously studied methods of estimating extreme quantiles:

(i) estimation based on the exponential distribution for exceedances over high thresholds,

(ii) estimation based on the Generalised Pareto distribution for

exceedances,

(iii) estimation based on the Gumbel distribution, and

(iv estimation based on the Generalised Extreme Value distribution.

Section 7 concerns the application of these results to theoretical comparisons among the procedures, in particular giving new results for the comparison of (i) and (iii). Finally, in Section 8 we give numerical results and comparisons with existing simulations.

2. EXTREME VALUE APPROXIMATIONS

The present section is a summary of the results of Smith (1988).

The starting point is the representation

$$-\log F(x) = \exp\left\{-\int_{x_*}^{x} \frac{dt}{\phi(t)}\right\}, \quad x_* < x < x^*, \tag{2.1}$$

where (x_*, x^*) is the range of the distribution. This is for the classical approach based on extreme value distributions; for the threshold approach we replace $-\log F(x)$ by $1-F(x)$ in (2.1). The two representations lead to the same results except for a few distributions, (e.g. uniform, exponential) for which the convergence is very rapid.

From (2.1) we deduce

$$\frac{-\log F(u+x\phi(u))}{-\log F(u)} = \exp\left\{-\int_0^x \frac{\phi(u)}{\phi(u+s\phi(u))} \, du\right\}$$

$$= \left\{1+x\phi'(y)\right\}^{-1/\phi'(y)} \tag{2.2}$$

for some y between u and $u+x\phi(u)$, by Taylor expansion and the mean value theorem. This assumes ϕ continuously differentiable.

Suppose now

$$\lim_{x \to x'} \phi'(x) = \gamma_o \quad (-\infty < \gamma_o < \infty) \tag{2.3}$$

This is essentially von Mises' condition for the existence of a limiting extreme value distribution. We now have three levels of approximation.

<u>First approximation</u> Replace $\phi'(y)$ in (2.2) by γ_o. Now if a_n, b_n are defined such that $-\log F(b_n) = n^{-1}$, $a_n = \phi(b_n)$, then replacing u by b_n in (2.2),

$$\lim_{n \to \infty} -n \log F(a_n x + b_n) = (1+\gamma_o x)^{-1/\gamma_o}$$

and hence

$$\lim_{n \to \infty} F^n(a_n x + b_n) = \exp\{-(1+\gamma_o x)^{-1/\gamma_o}\}, \tag{2.4}$$

valid wherever $1 + \gamma_o x > 0$. This is the <u>Generalised Extreme Value</u> distribution. In the special case $\gamma_o = 0$, the right hand side of (2.4) reduces to $\exp(-e^{-x})$, commonly called the <u>Gumbel</u> distribution.

<u>Second approximation</u> Replace $\phi'(y)$ in (2.2) by $\phi'(u)$. Then, defining a_n and b_n as above, and $\gamma_n = \phi'(b_n)$, we have the aproximation

$$F^n(a_n x + b_n) \approx \exp\left\{-(1+\gamma_n x)^{-1/\gamma_n}\right\} \tag{2.5}$$

which in general improves on (2.4). This is the <u>penultimate</u> approximation. In the case $\gamma_o = 0$, use of the penultimate approximation generally improves the rate of convergence (Cohen 1982, Gomes 1984). This is not so when $\gamma_o \neq 0$, but even here the quality of approximation for moderate n is generally improved by the penultimate approximation (Gomes and Pestana 1986, Smith 1988).

<u>Third approximation</u> The approximation may be further improved by expanding ϕ'. However, some care is needed to do this in a sufficiently general way. Simple Taylor expansion does not suffice. Smith (1988) proceeded under the assumption

$$\lim_{u\uparrow x^*} \frac{\phi'(u)\{\phi'(u+w\phi(u))-\phi'(u)\}}{g(u)h_\rho(1+w\phi'(u))} = c \tag{2.6}$$

for each w such that $1 + \gamma_0 w > 0$. Here ρ and c are real numbers, g is a non-negative function, and h_ρ is defined by

$$h_\rho(x) = \int_1^x v^{\rho-1}\,dv = \begin{cases} \dfrac{x^\rho-1}{\rho}, & \rho \neq 0, \\ \log x, & \rho = 0. \end{cases} \tag{2.7}$$

This assumption covers most common distributions. For instance:

(i) Assume $-\log F(x) = Cx^{-\alpha}(1 + Dx^{-\beta} + o(x^{-\beta}))$ as $x\to\infty$. This includes such cases as Pareto, Cauchy, t and F distributions. In this case $\gamma_0 = \alpha^{-1}$, $\rho = -\beta$, $g(u) = u^{-\beta}$, $c = -D\beta^2\alpha^{-3}(\beta-1)$.

(ii) Assume $x^* < \infty$ and

$$-\log F(x) = C(x^*-x)^\alpha\ \{1+D(x^*-x)^\beta + o((x^*-x)^\beta)\}$$

as $x \uparrow x^*$. Then $\gamma_0 = -\alpha^{-1}$, $\rho = \beta$, $g(u) = (x^*-u)^\beta$, $c = -D\beta^2\alpha^{-3}(\beta+1)$. This includes most distributions in the Weibull domain of attraction.

(iii) Assume $\gamma_0 = 0$ under what Cohen (1982) called class N, which includes most distributions in the Gumbel domain of attraction. In this case it is valid to make a Taylor expansion of ϕ. Equation (2.6) holds with arbitrary ρ, $g(u) = \phi(u)|\phi''(u)|$, $c = \pm 1$. This case covers the normal, lognormal and gamma distributions, amongst many others.

In all three cases, additional differentiability assumptions are involved, made precise in Smith (1988).

Substituting from (2.6) into (2.2), defining a_n, b_n, γ_n as before and $r_n = g(b_n)$, routine but tedious manipulations lead to

$$F^n(a_n x+b_n) \approx \exp[-(1+x\gamma_n)^{-1/\gamma_n}\{1+cr_n H_\rho(x,\gamma_n)\}] \tag{2.8}$$

whenever $1+x\gamma_n > 0$, where

$$H_\rho(x,\gamma) = \begin{cases} \dfrac{h_\rho(1+x\gamma)+\rho h_{-1}(1+x\gamma)-(\rho+1)\log(1+x\gamma)}{\rho(\rho+1)\gamma^3}, & 1+x\gamma>0 \\ 0, & 1+x\gamma\le 0 . \end{cases} \tag{2.9}$$

The rates of convergence for the three approximations are $O(\gamma_n - \gamma_o)$ for (2.4), $O(r_n)$ for (2.5) and $o(r_n)$ for (2.8). Under mild additional assumptions, these rates hold pointwise, uniformly over all x, in total variation norm, and finally in Hellinger distance. The Hellinger distance between two distributions with densities f and g is defined by

$$H(f,g) = [\int\{f^{\frac{1}{2}}(x) - g^{\frac{1}{2}}(x)\}^2 dx]^{\frac{1}{2}} .$$

The use of Hellinger distance in the context of extreme value approximations was first proposed by Reiss (1984) and, as will be seen in later sections, greatly simplifies the proofs of statistical results.

In the case of (2.8), there is a small problem in that the right hand side may not be a distribution function. In this case, however, a slight modification does converge in Hellinger distance at rate $o(r_n)$. The principal additional assumption for this is that $\gamma_o > -1/2$, an assumption which is natural in view of the regularity conditions for maximum

likelihood estimation (Smith 1985).

For the threshold approach, $-\log F$ is replaced by $1-F$ in (2.1) and (2.2). We then have three approximations for $\{1-F(u+x\phi(u))\}/\{1-F(u)\}$:

<u>First approximation</u> $(1+\gamma_o x)^{-1/\gamma_o}$, for $x>0$, $1+\gamma_o x>0$, $-\infty<\gamma_o<\infty$. This is the Generalised Pareto tail.

<u>Second approximation</u> $(1+\gamma_u x)^{-1/\gamma_u}$, where $\gamma_u = \phi'(u)$.

<u>Third approximation</u> $(1+\gamma_u x)^{-1/\gamma_u}\{1 + cr_u H_\rho(x,\gamma_u)\}$ (2.10)

where $r_u = g(u)$.

The errors of these approximations are, respectively, $O(\gamma_u-\gamma_o)$, $O(r_u)$, $o(r_u)$, and these are valid for Hellinger distance as before.

3. ESTIMATION BASED ON AN EXPONENTIAL TAIL

In this section we consider the estimation of an extreme quantile under the assumption that the conditional distribution of exceedances over a high threshold is exponential. By the previous section, this is the same as a Generalised Pareto approximation with $\gamma_o = 0$. The method is effectively the same as that of Weissman (1978).

The general framework, here and throughout the rest of the paper, is that we have N independent observations from a common unknown distribution function F, and we are interested in solving $1-F(q)=p$ for given small p. Fix a high threshold u, let k denote the (random) number of exceedances of

u, and let $Y_1,\ldots,Y_k$ denote the excesses over u. Given k, we have $Y_1,\ldots,Y_k$ independent with common distribution function

$$\frac{F(u+y)-F(u)}{1-F(u)} \approx 1-\exp\left\{-\frac{y}{\phi(u)}\right\}, \quad y > 0 \tag{3.1}$$

under the First Approximation. Here we assume $\gamma_0 = 0$.

The method that then suggests itself is to estimate $\theta = 1-F(u)$ by k/N and then fit an exponential distribution to $Y_1,\ldots,Y_k$, estimating ϕ by the sample mean $\bar{Y}$. These approximations lead to the estimator

$$\hat{q} = u + \bar{Y}\log\left[\frac{k}{Np}\right] \tag{3.2}$$

assuming q > u. Except for changes in notation, this is identical to equation (4.3) in Weissman (1978), except that Weissman took k rather than u as predetermined.

An approximation for the variance of $\hat{q}$ may be made by the usual delta method, assuming the exponential distribution of the excesses and an approximate Poisson distribution for k. This leads to

$$\mathrm{var}(\hat{q}) \approx \frac{\phi^2}{k}\log^2\left(\frac{k}{Np}\right) + \frac{\bar{Y}^2}{k} \approx \frac{\phi^2}{N\theta}(1+\eta^2) \tag{3.3}$$

where $\eta = \log(\theta/p)$.

The next step is to approximate the bias of $\hat{q}$ due to misspecification of the exponential distribution. In view of the results described in Section 2, the obvious thing to do is to use the next order of approximation, i.e. the Generalised Pareto distribution, in place of (3.1). In other words, we replace the right hand side of (3.1) by

$$1 - \left\{1 + \frac{\gamma y}{\phi(u)}\right\}^{-1/\gamma} \tag{3.4}$$

where $\gamma = \gamma_u = \phi'(u)$. Now, the mean of a Generalised Pareto distribution with parameters γ and ϕ is $(1-\gamma)^{-1}\phi$, assuming $\gamma<1$. Moreover, under the Generalised Pareto model the true quantile would be $u + \phi\gamma^{-1}(e^{\gamma\eta}-1)$. Hence

$$E(\hat{q}-q) = \phi\left\{\frac{\eta}{1-\gamma} - \frac{e^{\gamma\eta}-1}{\gamma} + o(\gamma)\right\}$$

$$= \phi\gamma\eta\{1-\eta/2 + o(1)\}\ . \tag{3.5}$$

Combining (3.3) and (3.5) leads to an approximation for mean squared error

$$MSE(\hat{q}) \approx \phi^2(u)\left[\frac{1+\eta^2}{N\theta} + \gamma^2\eta^2\left[1 - \frac{\eta}{2}\right]^2\right]. \tag{3.6}$$

So far, this development has been heuristic. I shall now outline a framework within which these formulae appear as rigorous asymptotic results. The intention of this is to clarify the status of the preceding approximations, as well as indicating a starting point for possible further theoretical development.

Suppose $N \to \infty$, $u = u_N \to x^*$, $\theta = \theta_N \to 0$, $N\theta_N \to \infty$, $\eta = \eta_N \to \eta_o$, where $0 < \eta_o < \infty$. Let $\phi_N = \phi(u_N)$, $\gamma_N = \phi'(u_N)$, $p_N = \theta_N\exp(-\eta_N)$. Suppose also

$$\gamma_N(N\theta_N)^{\frac{1}{2}} \to \delta \tag{3.7}$$

where δ is finite (possibly 0). Define

$$\psi_N^2 = \frac{\phi_N^2}{N\theta_N}(1 + \eta_N^{\ 2}). \tag{3.8}$$

Consider the following hypothetical model. For each N, k_N has a binomial distribution with parameters N, θ_N and, given $k_N = k$, the excesses $Y_1,\dots,Y_k$ are independent Generalised Pareto with parameters ϕ_N, γ_N.

4. ESTIMATION BASED ON A GENERALISED PARETO TAIL

The method of Section 3 will now be extended to one in which the exponential distribution (3.1) is replaced by the Generalised Pareto distribution for the purpose of constructing the estimator. This was the method studied in detail by Smith (1987), but the new probabilistic approximations of Section 2 allow the results to be developed in a much more coherent fashion.

Suppose, then, that (3.4) is used in place of the approximation in (3.1), and that estimators $\hat{\gamma}$, $\hat{\phi}$ are obtained for the parameters γ, ϕ. Throughout this paper it will be assumed that maximum likelihood is used as the method of estimation, though there is nothing to stop similar calculations being made for other estimators, such as the Probability Weighted Moments estimators of Hosking and Wallis (1987). As before, the exceedance probability θ is estimated by k/N, and this leads to a quantile estimate

$$\hat{q} = u + \hat{\phi}\hat{\gamma}^{-1}(\exp(\hat{\gamma}\,\hat{\eta}) - 1) \quad \text{where} \quad \hat{\eta} = \log\left[\frac{k}{Np}\right]. \tag{4.1}$$

In this section γ_o (from (2.3)) is arbitrary, but we assume $\gamma_o > -1/2$ for regularity.

Applying the delta method, with an approximate Poisson distribution for k and the covariance matrix of $(\hat{\gamma}, \hat{\phi})$ derived from the Fisher information matrix (Smith 1987), we obtain the approximation

$$\mathrm{var}(\hat{q}) \approx \frac{1+\gamma}{N\theta}\,(2\phi^2 D_1^2 - 2\phi D_1 D_2 + (1+\gamma)D_2^2) + \frac{\phi^2 e^{2\gamma\eta}}{N\theta} \tag{4.2}$$

where $\gamma = \gamma_u = \phi'(u)$, $\phi = \phi(u)$, $\eta = \log(\theta/p)$ and

$$D_1 = \frac{\partial q}{\partial \phi} = \frac{e^{\gamma\eta}-1}{\gamma}, \quad D_2 = \frac{\partial q}{\partial \gamma} = \phi\left[\frac{\eta e^{\gamma\eta}}{\gamma} - \frac{e^{\gamma\eta}-1}{\gamma^2}\right]. \tag{4.3}$$

Define $\hat{q}_N$ by (3.2) with suffices N added where appropriate. Also let q_N^* denote the exact quantile under the hypothetical model, i.e.

$$q_N^* = u_N + \phi_N(\exp(\gamma_N\eta_N) - 1)/\gamma_N .$$

For this model it is easy to make all the approximations rigorous and it follows that

$$P_N^*(\psi_N^{-1}(\hat{q}_N - q_N^*) \le x) \to \Phi\left[x - \delta\eta_0\left[1 - \frac{\eta_0}{2}\right]\left[1 + \eta_0^2\right]^{-\frac{1}{2}}\right] \qquad (3.9)$$

where Φ is the standard normal distribution function. Here P_N^* denotes the probability measure of the hypothetical model.

The hypothetical model differs from the true model only in the distribution of excesses over the threshold. By the results of Section 2, the Hellinger distance between the true and Generalised Pareto distributions is of $o(\gamma_N)$. Moreover, the Hellinger distance between two product measures grows in proportion to the square root of the number of components (Reiss 1984), so the Hellinger distance between P_N^* and the true probability measure, P_N say, is of $o(k_N^{\frac{1}{2}}\gamma_N)$. By (3.7), this tends to 0. Since Hellinger distance dominates total variance distance, it follows that $P_N^*(A_N) - P_N(A_N) \to 0$ for any sequence of events $\{A_N\}$. Hence (3.9) remains true with P_N replacing P_N^* and the true quantile q_N replacing q_N^*.

What this says, in effect, is that probability calculations based on (3.3), (3.5) and the normal distribution, are asymptotically correct. This avoids the direct question of whether (3.6) itself is valid as an asymptotic approximation, something which involves additional moment-convergence technicalities of the kind developed in Section 2.5 of Goldie and Smith (1987).

To approximate the bias of $\hat{q}$, the method is to use the higher -order approximation (2.10). Let us rewrite this in the form

$$\frac{F(u+y) - F(u)}{1-F(u)} \approx 1 - \left[1 + \frac{\gamma y}{\phi}\right]^{-1/\gamma}\left\{1 + \epsilon H_\rho\left[\frac{y}{\phi}, \gamma\right]\right\}, \tag{4.4}$$

where $\epsilon = \epsilon_u = cg(u)$, and take the right hand side of (4.4) to be exact over the range on which it is a valid distribution function.

The bias of $\hat{\phi}$ and $\hat{\gamma}$ are calculated by the method given in Section 2 of Smith (1987), which leads to the approximation

$$E\begin{Bmatrix}\hat{\phi} - \phi \\ \hat{\gamma} - \gamma\end{Bmatrix} \approx M^{-1} E\begin{Bmatrix}\dfrac{\partial \log g(Y)}{\partial \phi} \\ \dfrac{\partial \log g(Y)}{\partial \gamma}\end{Bmatrix} \tag{4.5}$$

where g is the Generalised Pareto density, M the Fisher information matrix, and the expected values are calculated under the model (4.4). Now

$$M^{-1} = (1+\gamma)\begin{bmatrix} 2\phi^2 & -\phi \\ -\phi & 1+\gamma \end{bmatrix},$$

$$\frac{\partial \log g(y)}{\partial \phi} = \frac{1}{\phi\gamma} - \frac{1}{\phi}\left[\frac{1}{\gamma} + 1\right]\left[1 + \frac{\gamma y}{\phi}\right]^{-1},$$

$$\frac{\partial \log g(y)}{\partial \gamma} = \frac{1}{\gamma^2}\log\left[1 + \frac{\gamma y}{\phi}\right] - \frac{1}{\gamma}\left[\frac{1}{\gamma} + 1\right]\left\{1 - \left[1 + \frac{\gamma y}{\phi}\right]^{-1}\right\}. \tag{4.6}$$

The range of y is $0<y<\infty$ if $\gamma \geq 0$, $0<y<-\gamma^{-1}$ if $\gamma<0$.

Suppose Y has distribution function (4.4). Define U to be $(1+\gamma Y/\phi)^{-1/\gamma}$. Now

$$E\left\{(1+\gamma Y/\phi)^r\right\} = E\left\{U^{-\gamma r}\right\} = 1+\gamma r\int_0^1 u^{-\gamma r-1}\, P\left\{U \leq u\right\} du .$$

Writing out the distribution function of U and using the identity

$$\int_0^1 u^{-\gamma r}\, h_\rho(u^{-\gamma})du = \frac{\gamma}{(1-\gamma r-\gamma\rho)(1-\gamma r)}$$

we deduce

$$E\left\{\left[1+\frac{\gamma Y}{\phi}\right]^r\right\} = \frac{1}{1-\gamma r} + \frac{\epsilon\gamma r}{(1-\gamma r-\gamma\rho)(1-\gamma r+\gamma)(1-\gamma r)^2} .$$

The conditions required for this are $\gamma r<1$, $\gamma r+\gamma\rho<1$. Taking the special cases $r=1$, $r=-1$ and the limit $r\to 0$, we deduce

$$E\left\{\frac{\partial \log g(Y)}{\partial\phi}\right\} = \frac{\epsilon}{\phi(1+\gamma-\gamma\rho)(1+\gamma)(1+2\gamma)} ,$$

$$E\left\{\frac{\partial \log g(Y)}{\partial\gamma}\right\} = \frac{\epsilon(3+2\gamma-2\gamma\rho)}{(1+\gamma-\gamma\rho)(1-\gamma\rho)(1+\gamma)(1+2\gamma)} .$$

Hence from (4.5) we have

$$E\{\hat\phi-\phi\} \approx -\frac{\epsilon\phi}{(1+\gamma-\gamma\rho)(1-\gamma\rho)} ,$$

$$E\{\hat\gamma-\gamma\} \approx \frac{\epsilon(2+\gamma-\gamma\rho)}{(1+\gamma-\gamma\rho)(1-\gamma\rho)} . \qquad (4.7)$$

Using (4.7) with a Taylor expansion in (4.1), we have

$$E(\hat q) \approx u + \phi\,\frac{e^{\gamma\eta}-1}{\gamma} + \epsilon\frac{(2+\gamma-\gamma\rho)D_2-\phi D_1}{(1+\gamma-\gamma\rho)(1-\gamma\rho)} .$$

Finally, for the true quantile under (4.4) we have

$$q = u + \phi\,\frac{e^{\gamma\eta}-1}{\gamma} + \epsilon\phi\, e^{\gamma\eta}\, H_\rho\left[\frac{e^{\gamma\eta}-1}{\gamma},\ \gamma\right] + o(\epsilon)$$

and hence

$$E(\hat q - q) \approx \epsilon\left[\frac{(2+\gamma-\gamma\rho)D_2-\phi D_1}{(1+\gamma-\gamma\rho)(1-\gamma\rho)} - \phi\, e^{\gamma\eta} H_\rho\left[\frac{e^{\gamma\eta}-1}{\gamma},\ \gamma\right]\right]$$

$$= \phi\,\epsilon\, B(\gamma,\rho,\eta) \qquad (4.8)$$

where

$$B(\gamma,\rho,\eta) = \frac{1}{(1+\gamma-\gamma\rho)(1-\gamma\rho)}\left\{(2+\gamma-\gamma\rho)\left[\frac{\eta e^{\gamma\eta}}{\gamma} - \frac{e^{\gamma\eta}-1}{\gamma^2}\right] - \frac{e^{\gamma\eta}-1}{\gamma}\right\}$$

$$- e^{\gamma\eta} H_\rho\left[\frac{e^{\gamma\eta}-1}{\gamma}, \gamma\right] . \tag{4.9}$$

Combining (4.2) and (4.8) one can obtain an approximation for the mean squared error. This can be used as a basis for deciding such questions as the optimal choice of threshold and whether, in the case of $\gamma_o=0$, the method of this section is preferable to the method of Section 2. Such questions are discussed at length in Smith (1987).

This may all be made rigorous by a similar process to that in Section 3. Making the same definitions as there, plus $\epsilon_N = cg(u_N)$, let ψ_N^2 denote the approximation to the variance of $\hat{q}_N$, obtained from (4.2) and (4.3). In place of (3.7) assume

$$\lim_{N\to\infty} \psi_N^{-1} \phi_N \epsilon_N B(\gamma_N,\rho,\eta_N) = \delta \tag{4.10}$$

where δ is finite. Since γ_N and η_N are converging to constants, this limit will exist whenever $k_N^{\frac{1}{2}} \epsilon_N$ converges. Consider a hypothetical model in which the approximation used in (4.4) is taken to be exact. Then all the preceding steps can be deduced by standard asymptotic arguments. The Hellinger distance between the hypothetical and true models tends to 0, leading to

$$P_N\left\{\psi_N^{-1} (\hat{q}_N - q_N) \le x\right\} \to \Phi\,(x-\delta) \tag{4.11}$$

where P_N is the true probability model. As before, this has the interpretation that probability calculations made on the basis of (4.2), (4.8) and the normal distribution, are asymptotically valid.

5. ESTIMATION BASED ON THE GUMBEL DISTRIBUTION

We now consider a classical extreme value procedure in which the sample of size N is divided into roughly k blocks of size n, and one of the classical extreme value distributions fitted to the block maxima. In this section we assume the Gumbel distribution

$$G(y) = \exp[-\exp\{-(y-\mu)/\sigma\}], \ -\infty < y < \infty \ . \tag{5.1}$$

This model only makes sense if it is the correct limiting distribution, i.e. if $\gamma_0=0$, and we assume that throughout this section.

If (5.1) is valid for $G = F^n$ and we are interested in the p-upper quantile of F, then

$$q = \mu - \sigma\log\ [-\log\{(1-p)^n\}] \ . \tag{5.2}$$

In limits as $n\to\infty$, $p\to0$, $np = Np/k \to e^{-\eta}$, (5.2) simplifies to $q = \mu+\sigma\eta$, which suggests the estimator

$$\hat{q} = \hat{\mu} + \hat{\sigma}\eta \ , \tag{5.3}$$

where $\hat{\mu}$, $\hat{\sigma}$ are the maximum likelihood estimates based on a Gumbel distribution for block maxima.

In the following Γ denotes the gamma function, $\psi = \Gamma'/\Gamma$ the digama function and ζ the Riemann zeta function. For $-\psi'(1) = 0.5772156649$ we write $\tilde{\gamma}$ (Euler's constant), in place of the more usual γ, to avoid confusion with the shape parameter. Also $\psi'(1) = \zeta(2) = \pi^2/6$, $\psi''(1) = -2\zeta(3) = -2.404113806$, $\psi'''(1) = 6\zeta(4) = \pi^4/15$ (Abramowitz and Stegun, 1964). The inverse Fisher information matrix is

$$M^{-1} = \frac{6\sigma^2}{\pi^2} \begin{bmatrix} \frac{\pi^2}{6} + (1-\tilde{\gamma})^2 & 1-\tilde{\gamma} \\ 1-\tilde{\gamma} & 1 \end{bmatrix} \tag{5.4}$$

(Gumbel 1958), and hence

$$\mathrm{var}(\hat{q}) \approx \frac{\sigma^2}{k} \left\{1 + \frac{6}{\pi^2} (1-\tilde{\gamma}+\eta)^2\right\}$$

$$= \frac{\sigma^2}{k} (1.10866 + 0.51404\eta + 0.60793\eta^2) \ . \tag{5.5}$$

To study the bias of $\hat{q}$, we embed the Gumbel distribution in the Generalised Extreme Value family

$$G(y) = \exp\left[-\left\{1 + \frac{\gamma(y-\mu)}{\sigma}\right\}^{-1/\gamma}\right] \tag{5.6}$$

where, for the approximation to F^n, we take $\mu = b_n$, $\sigma = a_n$, $\gamma = \gamma_n$ as in Section 2. The method is similar to (4.5), in that the expected values of the derivatives of the Gumbel log likelihood, with respect to μ and σ, are evaluated under (5.6).

If Y has distribution function (5.6) and density $g(y) = g(y;\mu,\sigma,\gamma)$, then we may write

$$Y = \mu + \frac{\sigma}{\gamma}(e^{\gamma Z}-1) = \mu + \sigma Z + \frac{\gamma\sigma Z^2}{2} + \cdots$$

where Z has a standard ($\mu = 0$, $\sigma = 1$) Gumbel distribution. Writing

$$\frac{\partial}{\partial\mu}\{\log g(Y;\mu,\sigma,0)\} = \frac{1}{\sigma}\left\{1-\exp\left[-\frac{Y-\mu}{\sigma}\right]\right\} = \frac{1}{\sigma}\left\{1-e^{-Z} + \frac{\gamma Z^2 e^{-Z}}{2} + O(\gamma^2)\right\} ,$$

$$\frac{\partial}{\partial\sigma}\{\log g(Y;\mu,\sigma,0)\} = \frac{1}{\sigma}\left[-1 + \frac{Y-\mu}{\sigma}\left\{1 - \exp\left[-\frac{Y-\mu}{\sigma}\right]\right\}\right]$$

$$= \frac{1}{\sigma}\{-1 + Z - Ze^{-Z} + \frac{\gamma}{2}(Z^2 - Z^2e^{-Z} + Z^3 e^{-Z}) + O(\gamma^2)\}$$

and using the formula

$$E\left\{Z^r e^{-sZ}\right\} = (-1)^r\, \Gamma^{(r)}(1+s) \,, \tag{5.7}$$

we find, to first order in γ,

$$E\left\{\frac{\partial}{\partial\mu}\log g(Y;\mu,\sigma,0)\right\} = \frac{\gamma}{2\sigma}\,\Gamma''(2) = \frac{\gamma}{2\sigma}\left[\frac{\pi^2}{6} - 2\tilde{\gamma} + \tilde{\gamma}^2\right]$$

$$= 0.41184\,\frac{\gamma}{\sigma} \,,$$

$$E\left\{\frac{\partial}{\partial\sigma}\log g(Y;\mu,\sigma,0)\right\} = \frac{\gamma}{2\sigma}\left\{\Gamma''(1) - \Gamma''(2) - \Gamma'''(2)\right\}$$

$$= \frac{\gamma}{2\sigma}\left\{2\tilde{\gamma} - \psi''(1) - 2 - \left[\frac{\pi^2}{2} - 3\right](1-\tilde{\gamma}) - (1-\tilde{\gamma})^3\right\}$$

$$= 0.33248\,\frac{\gamma}{\sigma} \,.$$

A calculation similar to (4.5) then leads to

$$E\begin{Bmatrix}\hat{\mu} - \mu \\ \hat{\sigma} - \sigma\end{Bmatrix} \approx \gamma\,\sigma \begin{bmatrix}0.54205 \\ 0.30798\end{bmatrix} . \tag{5.8}$$

As an aside, it should be pointed out that (5.8) is identical with Theorem 1 of Cohen (1987) when the mathematical technicalities of the latter are disregarded.

Application to (5.3), and comparison with the true quantile under the Generalised Extreme model, leads to the approximation

$$E(\hat{q} - q) \approx \gamma\,\sigma\,(0.54205 + 0.30798\eta - 0.5\eta^2) \,. \tag{5.9}$$

As in previous sections, these calculations may be made rigorous within a suitable asymptotic framework.

6. ESTIMATION BASED ON THE GENERALISED EXTREME VALUE DISTRIBUTION

The estimator is developed in the same way as in Section 5, but the Generalised Extreme Value distribution (5.6) is used in place of the Gumbel distribution (5.1). It is not necessary to assume $\gamma_0=0$, but we do take $\gamma_0 > -\frac{1}{2}$. The estimator is now taken to be

$$\hat{q} = \hat{\mu} + \hat{\sigma}\, \hat{\gamma}^{-1} \{\exp(\hat{\gamma}\, \hat{\eta})-1\} \tag{6.1}$$

with maximum likelihood estimates $\hat{\mu},\hat{\sigma},\hat{\gamma}$, fitted to the k block maxima.

The Fisher information matrix has been given by Prescott and Walden (1980) in the form

$$M = \begin{bmatrix} \frac{m_{11}}{\sigma^2} & \frac{m_{12}}{\sigma^2} & \frac{m_{13}}{\sigma} \\ \frac{m_{12}}{\sigma^2} & \frac{m_{22}}{\sigma^2} & \frac{m_{23}}{\sigma} \\ \frac{m_{13}}{\sigma} & \frac{m_{23}}{\sigma} & m_{33} \end{bmatrix} \tag{6.2}$$

where $m_{11} = P$, $m_{12} = \dfrac{\Gamma(2+\gamma)-P}{\gamma}$,

$$m_{13} = -\frac{1}{\gamma}\left[Q - \frac{P}{\gamma}\right], \quad m_{22} = \frac{1}{\gamma^2}\{1-2\Gamma(2+\gamma)+P\},$$

$$m_{23} = -\frac{1}{\gamma^2}\left[1 - \tilde{\gamma} + \frac{\{1-\Gamma(2+\gamma)\}}{\gamma} - Q + \frac{P}{\gamma}\right],$$

$$m_{33} = \frac{1}{\gamma^2}\left\{\frac{\pi^2}{6} + \left[1 - \tilde{\gamma} + \frac{1}{\gamma}\right]^2 - \frac{2Q}{\gamma} + \frac{P}{\gamma^2}\right\},$$

$$P = (1+\gamma)^2\, \Gamma(1+2\gamma),$$

$$Q = \Gamma(2+\gamma)\left\{\psi(1+\gamma) + \frac{1+\gamma}{\gamma}\right\}.$$

Here Γ and ψ are the gamma and digamma functions and $\tilde{\gamma}$ is Euler's constant. Hence we deduce

$$\mathrm{var}(\hat{q}) \sim \frac{\sigma^2}{k} \underset{\sim}{D}^T M_o^{-1} \underset{\sim}{D} \tag{6.3}$$

where M_o has entries (m_{ij}) (i.e. (6.2) with σ=1), $\underset{\sim}{D}^T = (D_1,D_2,D_3)$ and

$$D_1 = \frac{\partial q}{\partial \mu} = 1, \quad D_2 = \frac{\partial q}{\partial \sigma} = \frac{e^{\gamma\eta}-1}{\gamma},$$
$$D_3 = \frac{1}{\sigma}\frac{\partial q}{\partial \gamma} = \frac{\eta}{\gamma} e^{\gamma\eta} - \frac{e^{\gamma\eta}-1}{\gamma^2}. \tag{6.4}$$

The values of σ and γ are here taken to be those appropriate to the approximation of F^n, i.e. $\sigma = a_n$, $\gamma = \gamma_n$ in the notation of Section 2.

The bias of $\hat{q}$ will again be computed by embedding the Generalised Extreme Value distribution in a larger family, here (2.8). More precisely, consider the extended family

$$G(y) = \exp\left[-\left\{1 + \frac{\gamma(y-u)}{\sigma}\right]^{-1/\gamma}\left\{1 + \epsilon H_\rho\left[\frac{y-\mu}{\sigma}, \gamma\right]\right\}\right] \tag{6.5}$$

with five parameters $\mu,\sigma,\gamma,\rho,\epsilon$. This is used as an approximation to $F^n(y)$, with $\mu = b_n$, $\sigma = a_n$, $\gamma = \gamma_n$, $\epsilon = cg(b_n)$. As in Section 4, it suffices to define (6.5) over the range within which it is a valid distribution function, setting $G(y)$ to be 0 or 1 outside this range.

By analogy with (4.5), the biases of the parameter estimates are given by

$$E\begin{Bmatrix} \hat{\mu} - \hat{\mu} \\ \hat{\sigma} - \sigma \\ \hat{\gamma} - \gamma \end{Bmatrix} \approx M^{-1} E \begin{Bmatrix} \dfrac{\partial \log g(Y)}{\partial \mu} \\ \dfrac{\partial \log g(Y)}{\partial \sigma} \\ \dfrac{\partial \log g(Y)}{\partial \gamma} \end{Bmatrix}, \tag{6.6}$$

where M is the Fisher information (6.2), g is the Generalised Extreme Value density and Y is distributed according to (6.5). The main task is to evaluate these expectations in the limit as $\epsilon \to 0$.

First, note that

$$\frac{\partial \log g(y)}{\partial \mu} = \frac{1+\gamma}{\sigma}\left\{1 + \frac{\gamma(y-\mu)}{\sigma}\right\}^{-1} - \frac{1}{\sigma}\left\{1 + \frac{\gamma(y-\mu)}{\sigma}\right\}^{-1/\gamma-1},$$

$$\frac{\partial \log g(y)}{\partial \sigma} = \frac{1}{\sigma\gamma}\left[1 - \left\{1 + \frac{\gamma(y-\mu)}{\sigma}\right\}^{-1/\gamma}\right] - \frac{1}{\gamma}\frac{\partial \log g(y)}{\partial \mu},$$

$$\frac{\partial \log g(y)}{\partial \gamma} = -\frac{1}{\gamma} + \frac{1}{\gamma^2}\log\left\{1 + \frac{\gamma(y-\mu)}{\sigma}\right\}.$$

$$-\frac{1}{\gamma^2}\left\{1 + \frac{\gamma(y-\mu)}{\sigma}\right\}^{-1/\gamma}\log\left\{1 + \frac{\gamma(y-\mu)}{\sigma}\right\} - \frac{\sigma}{\gamma}\frac{\partial \log g(y)}{\partial \sigma}. \qquad (6.7)$$

Suppose a random variable Y has distribution function (6.5). Inverting (6.5), we may represent Y in the form

$$1 + \frac{\gamma(Y-\mu)}{\sigma} = e^{\gamma Z}\left\{1 + \gamma\,\epsilon\, H_\rho\left[\frac{e^{\gamma Z}-1}{\gamma}, \gamma\right] + O(\epsilon^2)\right\}$$

where Z has a reduced Gumbel distribution. Substituting in (6.7) and discarding terms of $O(\epsilon^2)$, we deduce

$$\frac{\partial \log g(Y)}{\partial \mu} = \frac{(1+\gamma)e^{-\gamma Z} - e^{-(1+\gamma)Z}}{\sigma}$$

$$-\frac{(1+\gamma)\epsilon H_\rho}{\sigma}\left\{\gamma e^{-\gamma Z} - e^{-(1+\gamma)Z}\right\},$$

$$\frac{\partial \log g(Y)}{\partial \sigma} + \frac{1}{\gamma}\frac{\partial \log g(Y)}{\partial \mu} = \frac{1}{\sigma\gamma}(1-e^{-Z}) + \frac{\epsilon\, e^{-Z} H_\rho}{\sigma\gamma},$$

$$\frac{\partial \log g(Y)}{\partial \gamma} + \frac{\sigma}{\gamma}\frac{\partial \log g(Y)}{\partial \sigma} = -\frac{1}{\gamma}(1-Z+Ze^{-Z}) + \frac{\epsilon H_\rho}{\gamma}\left\{1 + Ze^{-Z} - e^{-Z}\right\} \qquad (6.8)$$

where we have written H_ρ in place of $H_\rho\left[\frac{e^{\gamma Z}-1}{\gamma}, \gamma\right]$.

From the definition of H_ρ in (2.9) and using (5.7), we deduce

$$E\left\{Z^r e^{-sZ} H_\rho\left[\frac{e^{\gamma Z}-1}{\gamma}, \gamma\right]\right\} = \frac{(-1)^r}{\rho^2(\rho+1)\gamma^3}\left[\Gamma^{(r)}(1+s-\rho\gamma)\right.$$

$$\left. - \Gamma^{(r)}(1+s) + \rho^2\{\Gamma^{(r)}(1+s)-\Gamma^{(r)}(1+s+\gamma)\} + \rho(\rho+1)\gamma\,\Gamma^{(r+1)}(1+s)\right]. \tag{6.9}$$

When $\epsilon=0$, all the expressions in (6.8) of course have expectation 0, as is readily verified directly from (5.7). We therefore evaluate expectations of the $O(\epsilon)$ terms in (6.8), using (6.9), to deduce, as $\epsilon \to 0$,

$$E\left\{\frac{\partial \log g(Y)}{\partial \mu}\right\} \sim \frac{\epsilon}{\sigma} P(\gamma,\rho) ,$$

$$E\left\{\frac{\partial \log g(Y)}{\partial \sigma}\right\} \sim \frac{\epsilon}{\sigma} \left\{Q(\gamma,\rho) - \frac{1}{\gamma} P(\gamma,\rho)\right\} ,$$

$$E\left\{\frac{\partial \log g(Y)}{\partial \gamma}\right\} \sim \epsilon\left\{R(\gamma,\rho) - \frac{Q(\gamma,\rho)}{\gamma} + \frac{P(\gamma,\rho)}{\gamma^2}\right\} , \tag{6.10}$$

where

$$P(\gamma,\rho) = \frac{1+\gamma}{\rho^2(\rho+1)\gamma^3}\left[(1-\gamma\rho)\Gamma(1+\gamma-\gamma\rho) - \Gamma(1+\gamma)\right.$$

$$\left. + \rho^2\{\Gamma(1+\gamma)-(1+\gamma)\Gamma(1+2\gamma)\} + \rho(\rho+1)\gamma\{\Gamma(1+\gamma)+\Gamma'(1+\gamma)\}\right] ,$$

$$Q(\gamma,\rho) = \frac{1}{\rho^2(\rho+1)\gamma^4}\left[\Gamma(2-\gamma\rho)-1 + \rho^2\{1-\Gamma(2+\gamma)\} + \rho(\rho+1)\gamma(1-\tilde{\gamma})\right] ,$$

$$R(\gamma,\rho) = \frac{1}{\rho^2(\rho+1)\gamma^4}\left[\gamma\rho\{\Gamma(1-\gamma\rho) + \rho\Gamma(1+\gamma)-\rho-1\} - \Gamma'(2-\rho\gamma)\right.$$

$$\left. + 1-\tilde{\gamma}-\rho^2\{1-\tilde{\gamma}-\Gamma'(2+\gamma)\} - \rho(\rho+1)\gamma\left(\frac{\pi^2}{6} - 2\tilde{\gamma}+\tilde{\gamma}^2\right)\right] . \tag{6.11}$$

We have used the relations $\Gamma'(1) = -\tilde{\gamma}$, $\Gamma'(2) = 1-\tilde{\gamma}$, $\Gamma''(2) = \pi^2/6 - 2\tilde{\gamma}+\tilde{\gamma}^2$.

Equations (6.6) and (6.10) may be combined into

$$E \begin{bmatrix} \dfrac{\hat{\mu} - \mu}{\sigma} \\ \dfrac{\hat{\sigma} - \sigma}{\sigma} \\ \hat{\gamma} - \gamma \end{bmatrix} \sim \epsilon\, \underset{\sim}{B}(\gamma,\rho) \qquad (6.12)$$

where

$$\underset{\sim}{B}(\gamma,\rho) = M_o^{-1} \begin{bmatrix} P(\gamma,\rho) \\ Q(\gamma,\rho) - \dfrac{P(\gamma,\rho)}{\gamma} \\ R(\gamma,\rho) - \dfrac{Q(\gamma,\rho)}{\gamma} + \dfrac{P(\gamma,\rho)}{\gamma} \end{bmatrix},$$

M_o being as in (6.3).

We now apply these results to calculate the bias in $\hat{q}$. From (6.1) and (6.12) we deduce

$$E(\hat{q}) \approx \mu + \frac{\sigma e^{\gamma\eta}-1}{\gamma} + \epsilon\sigma\, \underset{\sim}{D}^T\, \underset{\sim}{B}(\gamma,\rho) \qquad (6.13)$$

where $\underset{\sim}{D}$ is as in (6.4). We also have, from the inverse of (6.5), that the true quantile under (6.5) satisfies

$$q \approx \mu + \sigma\frac{e^{\gamma\eta}-1}{\gamma} + \epsilon\sigma\, e^{\gamma\eta}\, H_\rho\left[\frac{e^{\gamma\eta}-1}{\gamma}, \gamma\right] \qquad (6.14)$$

Combining (6.13) and (6.14) we have

$$E(\hat{q}-q) \approx \epsilon\sigma\left\{\underset{\sim}{D}^T\, \underset{\sim}{B}(\gamma,\rho) - e^{\gamma\eta}\, H_\rho\left[\frac{e^{\gamma\eta}-1}{\gamma}, \gamma\right]\right\}. \qquad (6.15)$$

These aproximations may again be made rigorous using arguments similar to those in previous sections.

7. COMPARISONS OF THE METHODS

The results of the preceding four sections may be applied to a number of questions concerning comparisons among the methods. Among these questions are:-

1. Choice of threshold (threshold methods) or of block size (classical methods).

2. Choice between two-parameter and three-parameter approximations, assuming $\gamma_0=0$.

3. Choice between the threshold and classical approaches.

In each case, a meaningful comparison must take the bias as well as the variance term into account, as otherwise it would be possible to achieve very high accuracy by taking the threshold very low or the block size very small.

Problem 1 has been considered by Pickands (1975), Hall and Welsh (1985) and Smith (1987) in the threshold case, Cohen (1987, 1988) in the classical case. Problem 2 was also studied by Smith (1987) in the threshold case, Cohen (1988) in the classical case. Problem 3 has been considered by Cunnane (1973) and Rosbjerg (1985) in the hydrology literature, and in a preprint by J. Husler and J. Tiago de Oliveira, but these authors have considered only the variance of the estimators and have neglected the bias.

To illustrate how these ideas may be developed, I consider here the comparison of the methods of Section 3 and 5, assuming $\gamma_0= 0$.

Consider first the threshold model with exponential exceedances. Assume the total sample size N and the desired tail probability p are

fixed. Then equation (3.6) may be written in the form

$$\mathrm{MSE}(\hat{q}) \approx \phi^2 \left[\frac{1+\eta^2}{k} + \gamma^2\eta^2 \left[1 - \frac{\eta}{2}\right]^2\right] \tag{7.1}$$

where $k = N\theta$ denotes the expected number of exceedances of the threshold u and $\phi = \phi(u)$, $\gamma = \phi'(u)$, $\eta = -\log(Np/k)$.

Similarly, under the Gumbel model of Section 5, equations (5.5) and (5.9) lead to

$$\mathrm{MSE}(\hat{q}) \approx \phi^2 \left[\frac{1.10866 + 0.51404\eta + 0.60793\eta^2}{k} + \gamma^2(0.54205 + 0.30798\eta - 0.5\eta^2)^2\right] \tag{7.2}$$

where $\phi = \phi(b_n)$, $\gamma = \phi'(b_n)$, $\eta = -\log(Np/k)$, $n = N/k$ and b_n satisfies $-\log F(b_n) = n^{-1}$.

If k is fixed then u in (7.1) and b_n in (7.2) are (almost) the same; hence so are ϕ and γ in the two equations. Moreover, over the range of values of k which are of interest, both ϕ and γ vary only slightly, so we may effectively treat these as constants. (Precise justification of this last statement will not be made, but the key point is that $\phi(u+y\phi(u))/\phi(u)\to 1$, $\phi'(u+y\phi(u)/\phi'(u)\to 1$, as $u\uparrow x^*$ for fixed y. These properties follow from the assumptions made in Section 2).

It follows, then, that we may compare the two procedures by directly comparing the expressions (7.1) and (7.2), treating k as a free parameter. The comparison depends on N and p only through the product Np. As an example, Figure 1 shows the two mean squared errors plotted against k for $Np=1$, $\gamma=0.1$, $\phi=1$. The minimum values are 0.494 for the Gumbel procedure at $k=23$, 0.462 for the threshold procedure at $k=42$. Thus the optimal k is

almost twice as large for the threshold procedure as for the Gumbel procedure, and the ratio of minimum mean squared errors is 1.07 in favour of the threshold procedure. In comparative terms, very similar results were obtained for other values of Np and γ which were tried.

In the papers cited earlier, the comparison between the two procedures was based on variance alone, under the assumption that the same k is used in each. There is no reason, however, to make such an assumption. The present study thus favours the threshold method so long as k is chosen optimally (or nearly optimally), though in view of the asymptotic nature of the result and the fairly small differences between the procedures it would be wrong to read too much into this conclusion.

8. COMPUTATIONAL RESULTS.

This section describes numerical comparisons of all four procedures for a number of parent distributions.

Simulation studies of similar questions have previously been published by Boos (1984), Gomes (1986) and Joe (1987). Boos compared the exponential method of estimating extreme quantiles with the nonparametric quantile estimates, for a variety of values of k. Gomes made a comparison of the Gumbel and Generalised Extreme Value distributions for estimating extreme quantiles in the classical approach to extreme value theory, when the limiting (ultimate) approximation is Gumbel. Her results generally support the use of the Generalised Extreme Value distribution, especially for estimating the more extreme quantiles. Joe made a number of comparisons of bias and mean square error for all four procedures studied in this paper. Joe concluded that, in general, estimation based on the Generalised Pareto distribution is slightly superior to that based on the Generalised Extreme Value distribution, but that, in all cases it is important not to take k too large.

Some attempt that has been made to reproduce the results of these three papers using the approximations of Sections 3-6. My approximations support their broad conclusions but do not reproduce their detailed numerical results. This is probably because the sample sizes are too small for the asymptotic results to be reasonable. For instance, Joe assumed total sample size N=600 of which a typical run was based on k=30 blocks of size n=20. In the following discussion I assume a sample size N=10000. Although this is much larger than the sample sizes in the earlier studies it is not at all unreasonable for many applications in the hydrology/meteorology areas.

Figure 2 plots the root mean squared error for the estimator of the 1/10000-quantile, (i.e. $p=10^{-4}$), based on N=10000 observations, calculated using the approximations in Sections 3-6, for various parent distributions and across a wide range of k. Corresponding calculations were also made for other quantiles with generally similar results.

One property that is sometimes observed with these calculations is a "bias-cancellation" phenomenon - the bias becomes zero owing to a cancellation of the terms contributing to it. This is observed in Fig.2(a) which is based on the standard normal distribution for the observations from which the sample is drawn. The bias for the Generalised Pareto method is zero near k=2100 and that for the Generalised Extreme Value method near k=2300. This results in an unusual shape of the two curves, with the apparently optimal k very large. From a practical point of view it would be unwise to rely on being able to exploit the bias cancellation and the most significant feature of Fig.2(a) is that the Generalised Pareto method does better than the Generalised Extreme Value method for most of the range of k. The other two methods are not even shown because their mean squared errors are far larger. In contrast, Fig.2(b) shows the four plots for the lognormal distribution ($\sigma=1$). In this case the bias comparisons again favour the Generalised Pareto/Generalised Extreme Value procedures, but the variances are much lower for the exponential/Gumbel procedures, with the exponential coming off best. Fig.2(c) shows a gamma distribution (scale parameter 1, shape parameter 5) reflected about the origin - this comparison would also be valid for inference about the lower tail when the parent distribution is a three-parameter gamma. The exponential and Gumbel procedures are not applicable here because the reflected gamma distribution is not in the Gumbel domain of attraction. In this case the Generalised Pareto method

does better for large k, the comparison being rather similar to Fig.1. A comparison for the upper tail of a t_4 distribution (Fig.2(d)) leads to rather similar conclusions except for a rather drastic bias cancellation effect at the right hand side. Finally, the Weibull distribution function $1 - \exp\{-(x/\beta)^{\alpha}\}$ was tried, with α=0.5, β=0.1 in Fig.2(e), α=1.5, β=1 in Fig.2(f). Fig.2(e) shows the Gumbel and exponential procedures dominant, largely because the variances again dominate the comparison. In contrast, Fig.2(f) shows the Generalised Pareto/Generalised Extreme Value procedures dominant. In comparing these two Weibull distributions, it may well be important that the α=0.5 case is heavier-tailed than an exponential distribution and in this respect comparable with the lognormal, while the α=1.5 case is lighter-tailed than the exponential distribution and therefore comparable with the normal. The classification of distributions into lighter than exponential, approximately exponential and heavier than exponential tails was also made by Boos (1984), who referred to earlier unpublished work by Breiman, Stone and Gins.

Summarising the results so far, the following general observations may be made:

1. In cases in the domain of attraction of a Gumbel distribution the Generalised Pareto/Generalised Extreme Value procedures perform better than the exponential/Gumbel procedures when the tail is lighter than exponential, but the comparison appears to be reversed when the tail is heavier than exponential. Also in the "heavier than exponential" case the optimal k is much smaller.

2. In all cases the comparison between the threshold and corresponding classical procedure appears to be similar to Fig.1, i.e. the

classical procedure does better at small k but the threshold procedure achieves its minimum mean squared error at a larger value of k and is then superior. The exceptions to this are when bias cancellation is observed.

Gomes (1986) compared the Gumbel and Generalised Extreme Value methods by simulation, remarking that the Generalised Extreme Value method seems to do better as higher quantiles are estimated. However, both the distributions she used for simulation (normal, and a modified form of Weibull with $\alpha=4$) are lighter-tailed than exponential so the present study suggests her simulations were not extensive enough to support those conclusions. Boos (1984) also remarked on the difficulties of the heavier-than-exponential case, even suggesting that simple nonparametric estimators might do better in such cases. Another simulation study, though not directly treating the bias vs. variance aspect of the problem, led Hosking, Wallis and Wood (1985) to conclude that maximum likelihood estimation has poor sampling properties in the heavy-tailed ($\gamma>0$) case, and to propose the method of probability weighted moments as an alternative. All these studies point to the need for a more detailed theoretical study of the heavy-tailed case, including perhaps the development of alternative estimators with smaller mean squared errors than maximum likelihood.

So far no indication has been given of the accuracy of the proposed approximations to the bias. A theoretical way to assess this is as follows. Suppose we have a very large sample from F^n for a given finite n, or from a threshold distribution for fixed threshold. The parameters of the fitted model (respectively, Generalised Extreme or Generalised Pareto) will converge to those values which maximise the expected log likelihood of the fitted model under the true distribution. These

limiting parameters may be calculated by a combination of numerical integration and numerical optimisation, and result in exact expressions for the bias - "exact" in the sense that they do not rely on any approximations for the distribution function. (They are still only approximations for finite k.) In Table 1 this is calculated for n=100 or 500 and the distributions used in Figure 2. Three approximations to γ are shown: the crude approximation $\phi'(u)$ of Section 2, the "bias corrected" approximation obtained from (4.7) or (6.12) and the "exact" value given by the procedure just described. Also shown are biases for two quantile estimates where, for N=10000, q_1 corresponds to p=.0001 (as in Figure 2) and q_2 to p=.0005. These are expressed as a percentage of the absolute value of the quantity being estimated. These results show that, although in most cases our approximations are the right order of magnitude, one would have to go to even larger samples before they were really accurate. In one sense this does not matter, because in practice we would not have the information to obtain the exact biases anyway, so a more important feature of our results is that they lead to the right qualitative conclusions.

9. SUMMARY AND CONCLUSIONS

The main purpose of the paper has been to obtain approximate expressions for the bias and variance of four established extreme-value procedures, and to use these to study the optimal value of k and to compare the four procedures in terms of mean squared error. The studies have in general supported the qualitative conclusions of simulation studies by other authors and have also suggested some new aspects. In particular, threshold procedures tend to require a larger optimal k than classical procedures, and then to achieve a lower mean squared error. Concerning the comparison between exponential/Gumbel procedures on the one hand, and Generalised Pareto/ Generalised Extreme Value procedures on the other, in many cases the study supports the latter, but not in heavy-tailed cases in the Gumbel domain of attraction (lognormal, Weibull with $\alpha<1$).

Important open questions are:

1. Data-based (adaptive) estimation of k. This may be best achieved by trying to estimate the bias, and then choosing k to minimise the estimated mean squared error. It seems to me that "general" methods of bias estimation, such as the jackknife and bootstrap, are unlikely to work in their usual form, in view of the rather specialised nature of the bias problem, but some modifications to take account of this may be possible. An asymptotically efficient, but highly artificial, proposal along these lines was made by Hall and Welsh (1985), in a more restricted setting than the one considered here. The Hall-Welsh result does serve to indicate what is theoretically possible.

2. Are there better estimation procedures than maximum likelihood? The well-known asymptotic optimality of maximum likelihood may not apply when

bias is taken into account, and indeed Csorgo, Deheuvels and Mason (1985) have indicated one way to improve it, though again in the more restricted setting of estimating the Pareto index. However, there may be simpler estimates than theirs which would have better properties than maximum likelihood, the probability-weighted moments estimator being presumably a candidate. Theoretical investigation of other estimators is possible along the same lines as developed for maximum likelihood estimators in this paper.

3. Another possibility is to expand the class of model distributions to something wider than the Generalised Extreme Value of Generalised Pareto Classes. There is considerable discussion along these lines in parts of the hydrology literature, the theoretical status of which is ill-defined, but there are obvious possibilities such as combining extreme value distributions with a family of transformations. Since this would inevitably involve increasing the variance of the estimators (because of the extra parameters), the only theoretical way to assess the idea would be in terms of some form of trade-off between bias and variance.

Acknowledgement

The paper was started at the Center for Stochastic Processes, University of North Carolina, and finished at the Department of Statistics, University of British Columbia. I would like to thank Ross Leadbetter and Harry Joe, respectively, for the invitations. The work was supported by the U.S. Air Force of Scientific Research, grant number F 49620 85C 0144 and by the Natural Sciences and Engineering Research Council of Canada.

TABLE 1 : EXACT BIAS CALCULATIONS

PARENT DIST.	METHOD	N	ESTIMATED γ			PERCENT BIAS IN q_1		PERCENT BIAS IN q_2	
			CRUDE	BC	EXACT	APPROX	EXACT	APPROX	EXACT
Normal	GEV	100	−.127	−.095	−.101	− .17	− .16	.40	.26
Normal	GPD	100	−.127	−.074	−.089	.16	.06	.31	.20
Lognormal (1)	GEV	100	.248	.224	.239	3.15	2.11	− .60	.03
Lognormal (1)	GPD	100	.248	.200	.222	.56	.43	−1.10	−.31
Refl.Gamma (5)	GEV	100	−.315	−.275	−.266	1.21	3.93	1.60	2.38
Refl.Gamma (5)	GPD	100	−.315	−.267	−.258	1.21	2.06	.85	1.09
t_4	GEV	100	.196	.212	.211	−2.08	−1.87	− .03	−.09
t_4	GPD	100	.196	.231	.225	− .46	− .50	.27	.18
Weibull (0.5,0.1)	GEV	500	.161	.129	.145	− .74	− .10	− .70	−.41
Weibull (0.5,0.1)	GPD	500	.161	.099	.127	−1.11	− .33	− .14	−.08
Weibull (1.5,1.0)	GEV	100	−.072	−.055	−.057	− .17	− .13	.22	.11
Weibull (1.5,1.0)	GPD	100	−.072	−.043	−.053	.07	.00	.18	.10

REFERENCES

Abramowitz, M. and Stegun, I.A. (1964), Handbook of Mathematical Functions. National Bureau of Standards, Washington D.C., reprinted by Dover, New York.

Beirlant, J. and Teugels, J. (1987). The asymptotic behaviour of Hill's estimator. Theory Probab. Appl. 31, 463-469.

Boos, D.D. (1984), Using extreme value theory to estimate large percentiles. Technometrics 26, 33-39.

Cohen, J.P. (1982), Convergence rates for the ultimate and penultimate approximations in extreme value theory, Adv. Appl. Prob. 14, 833-854.

Cohen, J.P. (1986), Large sample theory for fitting an approximate Gumbel model to maxima. Sankhya A 48, 372-392.

Cohen, J.P. (1988), Fitting extreme value distribution to maxima. University of Kentucky.

Csörgő, M., Horvath, L. and Revesz, P. (1987), On the optimality of estimating the tail index and a naive estimator. Australian J. Statist. 29, 166-178.

Cunnane, C. (1973), A particular comparison of annual maxima and partial duration series methods of flood frequency prediction. J. Hydrology 18, 257-271.

Davis, R. and Resnick, S. (1984), Tail estimates motivated by extreme value theory. Ann. Statist, 12, 1467-1487.

Davison, A.C. (1984), Modelling excesses over high thresholds, with an application. In Statistical Extremes and Applications, J. Tiago de Oliveira (ed.), Reidel, Dordrecht, 461-482.

Goldie, C.M. and Smith, R.L. (1987), Slow variation with remainder: Theory and Application. Quart. J. Math. Oxford (2) 38, 45-71.

Gomes, M.I. (1984), Penultimate limiting forms in extreme value theory. Ann. Inst. Statist. Math. 36, 71-85.

Gomes, M.I. (1986), Penultimate versus ultimate in statistical theory of extremes, a simulation study. Comp. Stat. Data Anal. 4, 257-267.

Gomes, M.I. and Pestana, D.D. (1987). Nonstandard domains of attraction and rates of convergence. In, M.L. Puri, J.P. Vilaplana and W. Wertz (eds.), New Perspectives in Theoretical and Applied Statistics, 467-477. John Wiley.

Gumbel, E.J. (1958), Statistics of Extremes. Columbia University Press, New York.

Hall, P. (1982), On some simple estimates of an exponent of regular variation. J.R. Statist. Soc. B 44, 37-42.

Hall, P. and Welsh, A.H. (1984), Best attainable rates of convergence for estimates of parameters of regular variation. Ann. Statist. 12, 1079-1084.

Hall, P. and Welsh, A.H. (1985), Adaptive estimates of parameters of regular variation. Ann. Statist. 13, 331-341.

Hosking, J.P.M. (1984), Testing whether the shape parameter is zero in the generalized extreme-value distribution. Biometrika 71, 367-374.

Hosking, J.R.M., and Wallis, J.R. (1987), Parameter and quantile estimation for the generalized Pareto distribution. Technometrics 29, 339-349.

Joe, H. (1987), Estimation of quantiles of the maximum of N observations. Biometrika 74, 347-354.

N.E.R.C. (1975), The Flood Studies Report. The Natural Environment Research Council, London.

Pickands, J. (1975), Statistical inference using extreme order statistics. Ann. Statist. 3, 119-131.

Prescott, P. and Walden, A.T. (1980), Maximum likelihood estimation of the parameters of the generalized extreme value distribution. Biometrika 67, 723-724.

Reiss, R.-D. (1984), Statistical inference using appropriate extreme value models. Preprint 124, University of Siegen.

Reiss, R.D. (1987), Estimating the tail index of the claim size distribution. Deutsche Gesellschaft fur Versicherungsmathematik 18, 21-25.

Rosbjerg, D. (1985), Estimation in partial duration series with independent and dependent peak values. J. Hydrology 76, 183-195.

Smith, R.L. (1984), Threshold method for sample extremes. In Statistical Extremes and Applications, J. Tiago de Oliveira (ed.), 621-638, Reidel, Dordrecht.

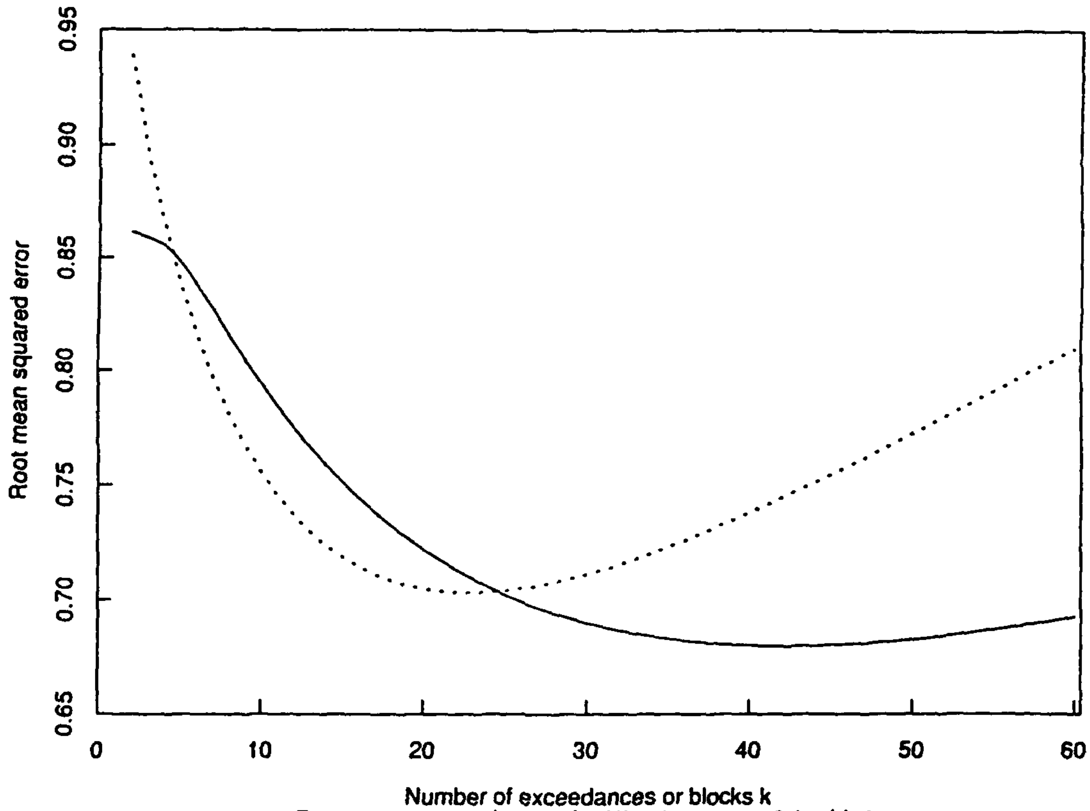
Smith, R.L. (1985), Maximum likelihood estimation in a class of non-regular cases. Biometrika 72, 67-92.

Smith, R.L. (1987), Estimating tails of probability distributions. Ann. Statist. 15, 1174-1207.

Smith, R.L. (1988), Approximations in extreme value theory. Submitted.

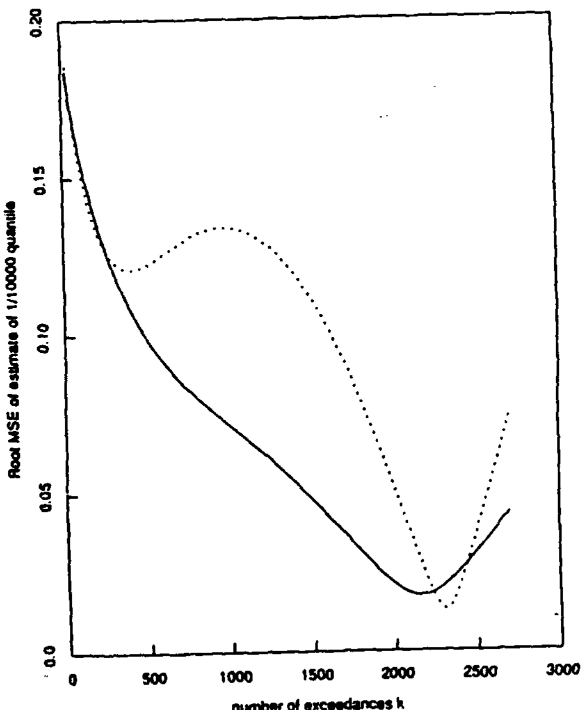
Weissman, I. (1978), Estimation of parameters and large quantiles based on the k largest observations. J. Amer. Statist. Assoc. 73, 812-815.

FIG. 1: COMPARISON OF EXPONENTIAL & GUMBEL PROCEDURES



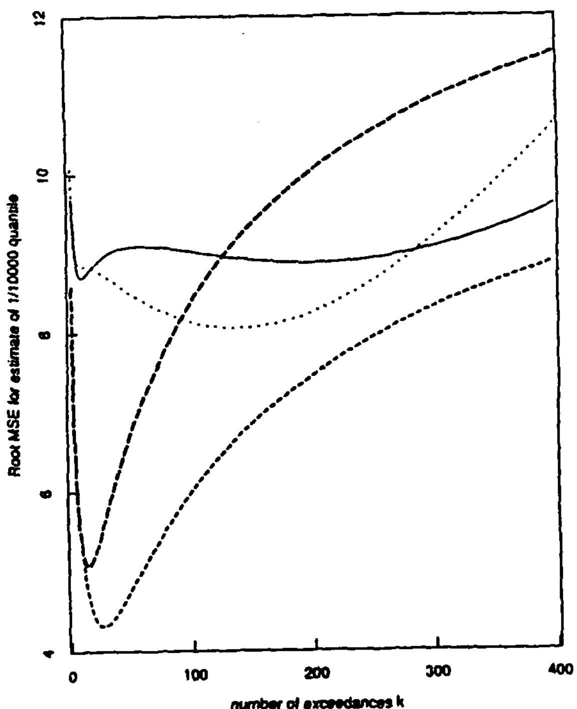
Root mean squared errors for N*p=1, gamma=0.1, phi=1,
Exponential model (continuous line), Gumbel model (dotted)

FIG. 2(a): 10000 OBSNS., STANDARD NORMAL



GPD (continuous line) and GEV (dotted line)

FIG. 2(b): 10000 OBSNS., LOGNOR (SIGMA 1)



GPD (cont.), GEV (dotted), Exp. (short lines), Gum. (long lines)

FIG. 2(c): 10000 OBSNS., REFL. GAMMA(5)

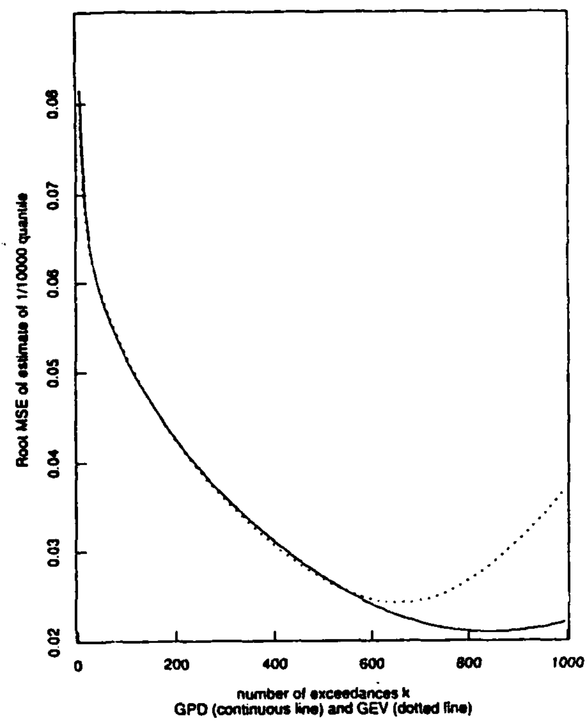


FIG. 2(d): 10000 OBSNS., t(4) DIST.

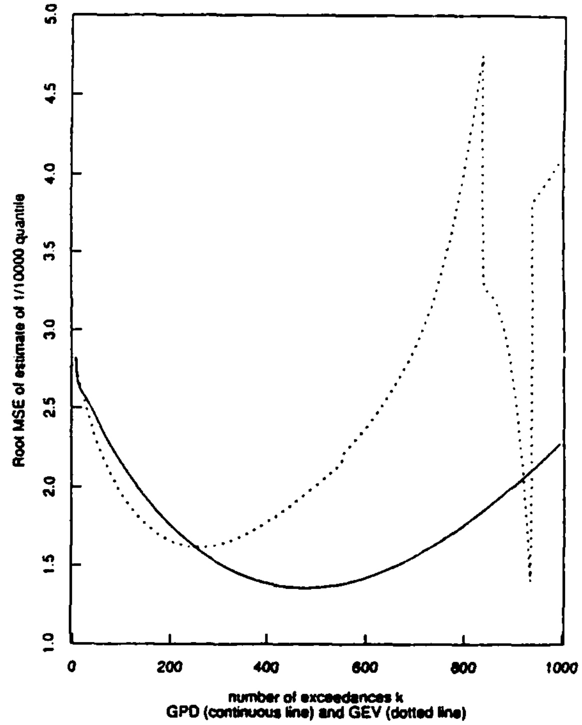


FIG. 2(e): 10000 OBSNS., WEIBULL (0.5,0.1)

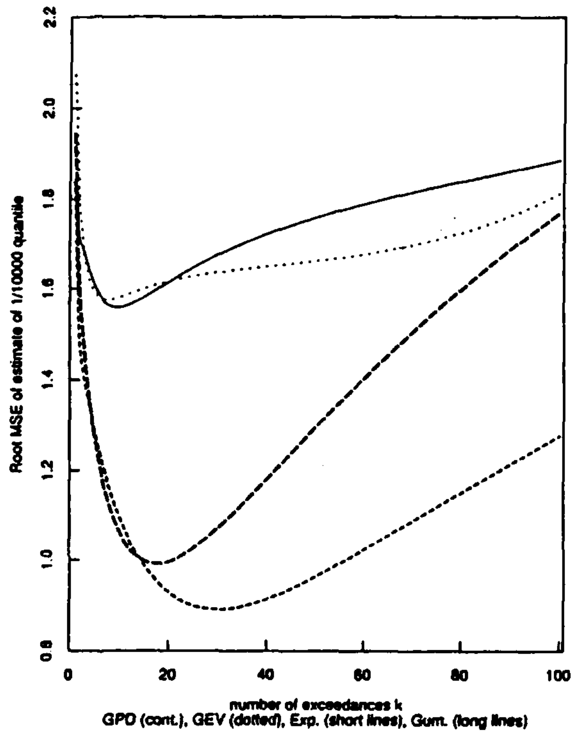
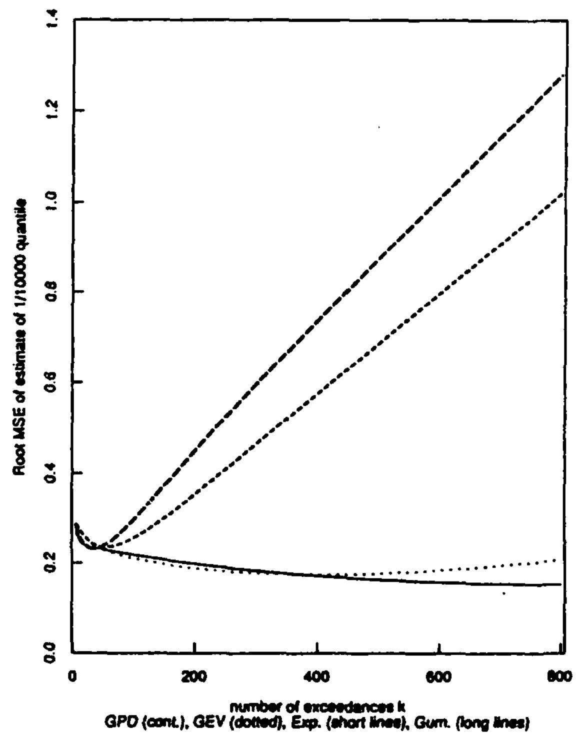


FIG. 2(f): 10000 OBSNS., WEIBULL (1.5,1.0)



222. I. Mitoma, Weak solution of the Langevin equation on a generalized functional space, Feb. 88.

223. L. de Haan, S.I. Resnick, H. Rootzén and C. de Vries, Extremal behaviour of solutions to a stochastic difference equation with applications to arch-processes, Feb. 88.

224. O. Kallenberg and J. Szulga, Multiple integration with respect to Poisson and Lévy processes, Feb. 88. *Prob. Theor. Rel. Fields*, 1989, to appear.

225. D.A. Dawson and L.G. Gorostiza, Generalized solutions of a class of nuclear space valued stochastic evolution equations, Feb. 88.

226. G. Samorodnitsky and J. Szulga, An asymptotic evaluation of the tail of a multiple symmetric α-stable integral, Feb. 88. *Ann. Probability*, to appear.

227. J.J. Hunter, The computation of stationary distributions of Markov chains through perturbations, Mar. 88.

228. H.C. Ho and T.C. Sun, Limiting distribution of nonlinear vector functions of stationary Gaussian processes, Mar. 88.

229. R. Brigola, On functional estimates for ill-posed linear problems, Apr. 88.

230. M.R. Leadbetter and S. Nandagopalan, On exceedance point processes for stationary sequences under mild oscillation restrictions, Apr. 88. *Oberwolfach Conf. on Extremal Value Theory*, J. Hüsler and R. Reiss, eds., Springer, to appear.

231. S. Cambanis, J. P. Nolan and J. Rosinski, On the oscillation of infinitely divisible processes, Apr. 88.

232. G. Hardy, G. Kallianpur and S. Ramasubramanian, A nuclear space-valued stochastic differential equation driven by Poisson random measures, Apr. 88.

233. D.J. Daley, T. Rolski, Light traffic approximations in queues (II), May 88.

234. G. Kallianpur, I. Mitoma, R.L. Wolpert, Diffusion equations in duals of nuclear spaces, July 88.

235. S. Cambanis, Admissible translates of stable processes: A survey and some new models, July 88.

236. E. Platen, On a wide range exclusion process in random medium with local jump intensity, Aug. 88.

237. R.L. Smith, A counterexample concerning the extremal index, Aug. 88.

238. G. Kallianpur and I. Mitoma, A Langevin-type stochastic differential equation on a space of generalized functionals, Aug. 88.

239. C. Houdré, Harmonizability, V-boundedness, (2,P)-boundedness of stochastic processes, Aug. 88.

240. G. Kallianpur, Some remarks on Hu and Meyer's paper and infinite dimensional calculus on finitely additive canonical Hilbert space, Sept. 88.

241. L. de Haan, A Brownian bridge connected with extreme values, Sept. 88.

242. O. Kallenberg, Exchangeable random measures in the plane, Sept. 88.

243. E. Masry and S. Cambanis, Trapezoidal Monte Carlo integration, Sept. 88.

244. L. Pitt, On a problem of H.P. McKean, Sept. 88.

245. C. Houdré, On the prediction theory of multivariate (2,p)-bounded processes, Sept. 88.

246. C. Houdré, Fourier integrals for stochastic processes and dilation of vector measures, Sept. 88.

247. J. Mijnheer, On the rate of convergence in Strassen's functional law of the iterated logarithm, Sept. 88.

248. G. Kallianpur and V. Perez-Abreu, Weak convergence of solutions of stochastic evolution equations on nuclear spaces, Oct. 88.